DIRECT PHASING IN CRYSTALLOGRAPHY : STATISTICAL APPROACH WITH MULTIMINIMA SCORE FUNCTIONS

By Alexandre Urzhumtsev
(sacha@lcm3b.uhp-nancy.fr)
LCM3B, Universite Henri Poincare, Nancy 1, France

Many structure determination techniques are based on the search for the global minimum of some score function. Our study of different criteria revealed that such approach is not relevant to the low resolution *ab initio* phasing in macromolecular crystallography. For all "natural" restraints on electron density or structure factors tried in our works [1-10] the best values of the score function were usually attained for phase sets significantly different from the correct phases while best found phase sets resulted often in intermediate values of the score function.

Nevertheless, the considered score functions produce a statistical improvement of the available phase sets, for example, generated randomly. A percentage of phase sets reasonably close to the exact values is higher for the ensemble of the selected phase sets independently on their percentage in the random population. Such 'enrichment' by reasonably good sets allows in general to obtain a searched crystallographic image through a simple averaging of the selected variants.

A more advanced statistical *ab initio* phasing approach has been developed on the base of this analysis. The search of the solution can be done either directly in the phase space (where every point is a phase set for given structure factors) or in the multidimensional space of parameters (every set of which allows to calculate a corresponding phase set; see for example [2-4]). The procedure consists in several steps :

- generation of a random population of points in the search space and selection of those points which resulted in a reasonable value of the score function;
- cluster analysis of the selected population in order to identify compact groups of the selected variants;
- averaging variants inside every identified cluster.

The procedure provides with a few alternative solutions (usually, a single solution) of the problem considered, which are investigated in the following tests if the solutions are several.

In one of applications, the first crystallographic image of the ribosomal particle 50S has been found *ab initio* [3]. Recently, an application of the developed procedures to the low resolution *ab initio* study of Low Density Lipoprotein complex [9] has been reported. Some latest results show that with the proper choice of the search parameters and the score function, the image resolution at least of 4-5 Å can be reached [10] starting from the low resolution end.

[1] Lunin, V., Urzhumtsev, A., Skovoroda, T. (1990) *Acta Cryst.*, **A46**, 540-544
[2] Lunin, V., Lunina, N., Petrova, T., Vernoslova, E., Urzhumtsev, A., Podjarny, A. (1995) *Acta Cryst.*, **D51,** 896-903
[3] Urzhumtsev, A., Vernoslova, E., Podjarny, A. (1996) *Acta Cryst.*, **D52**, 1092-1097
[4] Lunin, V., Lunina, N., Petrova, T., Urzhumtsev, A., Podjarny, A. (1998) *Acta Cryst.*, **D54,** 726-734
[5] Lunin, V., Lunina, N., Urzhumtsev, A. (2000) *Acta Cryst.*, **A56,** 375-382
[6] Lunin, V., Lunina, N., Petrova, T., Skovoroda, T., Urzhumtsev, A., Podjarny, A. (2000) *Acta Cryst.*, **D56,** 1223-1232
[7] Urzhumtsev, A., Lunina, N., Skovoroda, T., Podjarny, A., Lunin, V. (2000) *Acta Cryst.*, **D56,** 1233-1244
[8] Petrova, T., Lunin, V., Podjarny, A. (2000) *Acta Cryst.*, **D56,** 1245-1252

[9] Lunin, V., Lunina, N., Ritter, S., Frey, I., Keul, J., Diederichs, K., Podjarny, A., Urzhumtsev, A., Baumstark, M. (2000) *Acta Cryst.*, **D56,** in press
[10] Lunina, N., Lunin, V., Urzhumtsev, A. (2000) *ECM-19 Abstracts, XIXth European Cryst.Meeting, 25-31 August 2000, Nancy, France*, 62